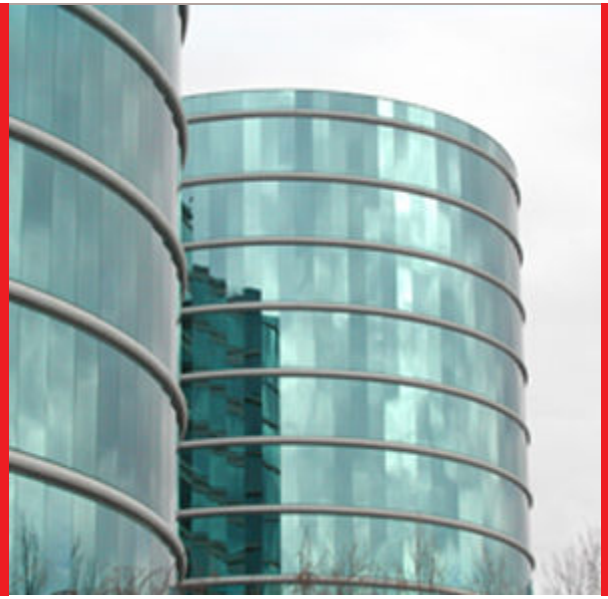


ORACLE

Oracle Database 11g XML DB

Mark Drake

Manager, Product Management

XML Technology, Oracle Corporation

Agenda

- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
- Customers and Partners

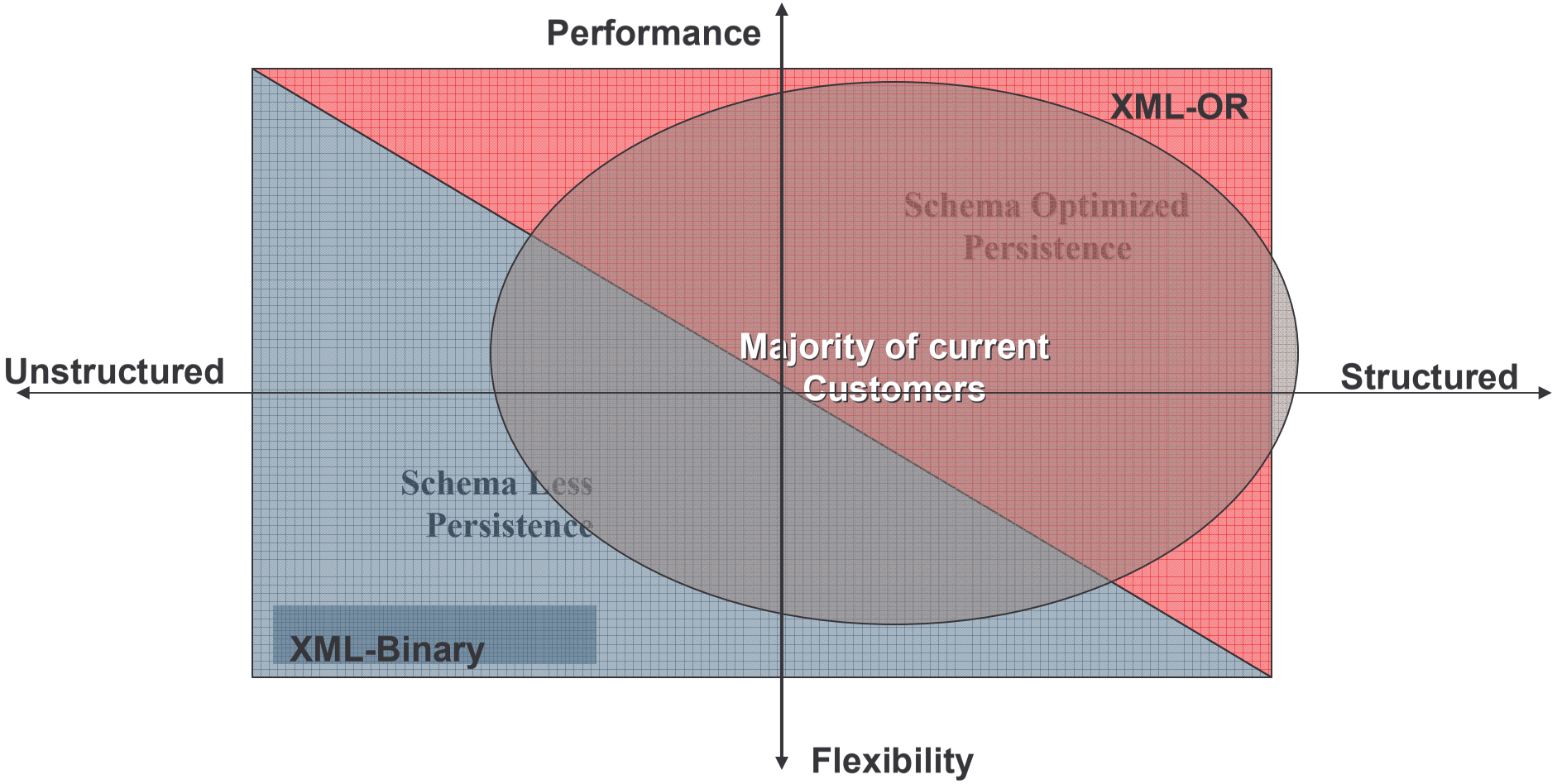
Why XML ?

- Common paradigm for managing structured and unstructured data
  - Flexible data model equally suited to data-centric and document centric application development
- Open, vendor-neutral standards, driven by W3C
  - XML, XMLSchema, XQuery, SOAP
- Widely adopted in multiple industries
  - Growth of XML based standards
    - XBRL, FPML, HL7, GJXML
- Primary storage format for next generation office productivity suites

Why XML in the database ?

- XML being used to manage mission critical information
  - Interchange with external organizations
  - Web Services
- Need to manage XML effective and efficiently
  - Number and size of documents increasing
  - Reliability, Scalability, Availability
  - Security
  - Compliance
- Accurate and fast information location and retrieval

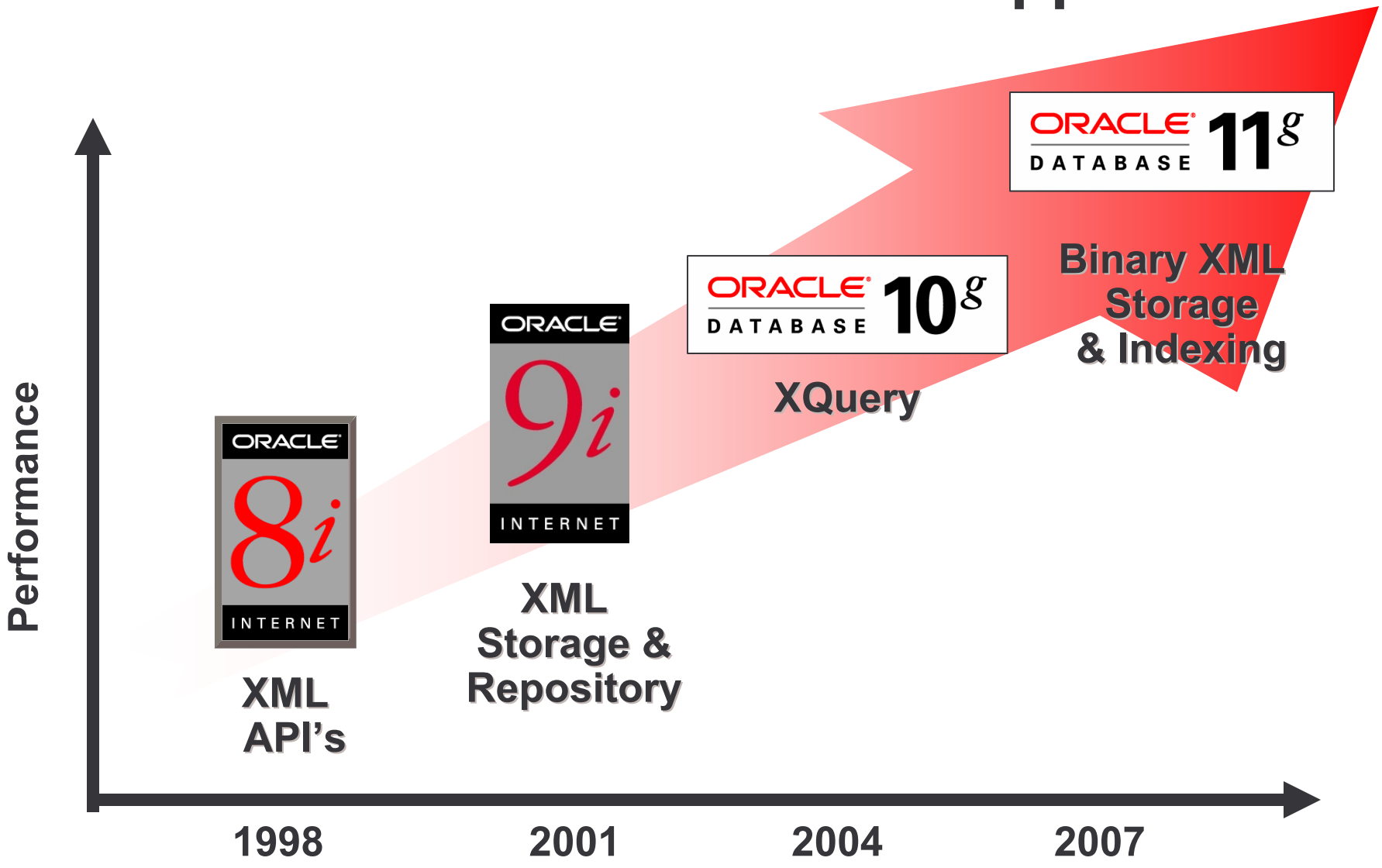
XML Use Cases



Agenda

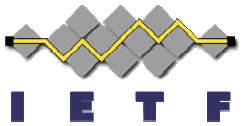
- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
- Customers and Partners

Evolution of Oracle's XML Support

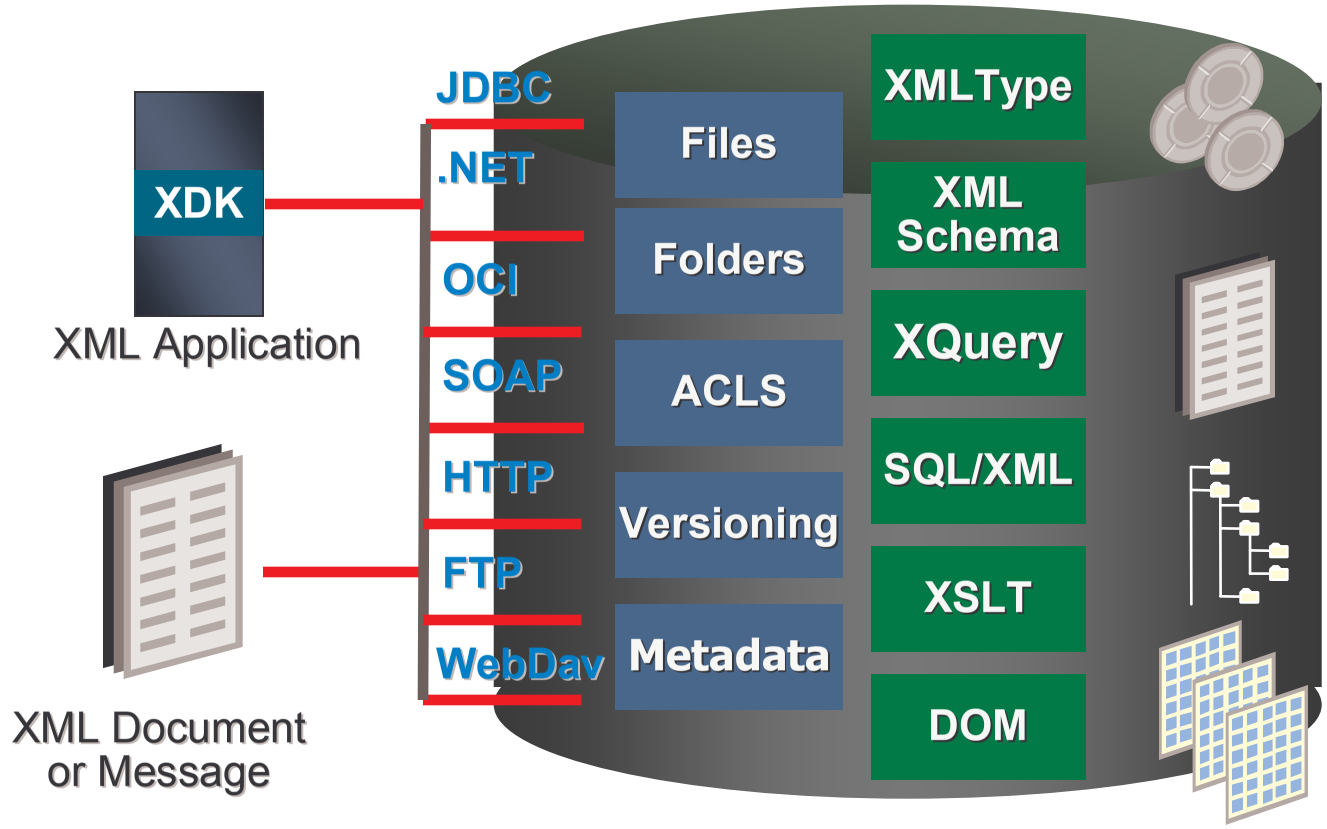


What is Oracle XML DB ?

- Native, flexible, high-performance, scaleable XML storage and processing
- XML repository allowing document centric development and access
- Implementations of key W3C standards
  - XML, Namespaces, XPath, XML Schema, XSLT, WebDAV
  - First major implementation of XQuery specification
- SQL/XML:2003 and 2006 compliant XML publishing
- SQL/XML and XML/SQL interoperability.



Oracle XML DB Overview



Schema-Based XML

- Object-Relational Storage
- Binary Storage

Schema-Less XML

- Binary Storage
- Text Storage

Relational Content

- XML Views

XML Indexing

- B-Tree Indexing
- XML Indexing
- Functional Indexing
- Full-Text Indexing

Access Paradigms

- Table / Row
- File / Folder
- Content and Meta-data

XML DB value propositions

- Fast and easy native XML application development
- Hybrid database
  - SQL centric access to XML content
  - XML centric access to relational content
- Multiple XML storage options allow tuning for optimal application performance
  - Application code is totally independent of storage model
  - Optimized storage and indexing for structured and unstructured XML
- XML DB repository enables document centric integrity and security models

XML DB Customers

- SQL Centric
  - WMG, Thyssen, UPS
  - Temenos
- XML Centric
  - Energysys/BP, Motorola
- Document Centric
  - State of California LDC
  - Nextance

Agenda

- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
- Customers and Partners

What's new in Oracle Database 11g ?

- Increased Flexibility for Schema-optimized storage
  - Zero down-time for schema change
  - Partitioning support
  - Intelligent defaults
  - General / Performance Enhancements
- Better management and performance for Schema-Less XML scenarios
  - Binary XML
  - XML Index
- Direct integration in to Web Service Architectures
- Programmable Repository and Compound Documents
- Improved Security
- XDK Enhancements

Agenda

- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
  - Improvements for Schema-Optimized XML
  - Schema-Less XML
  - XML Index
  - Database-Native Web Services
  - Repository Improvements
  - Security Enhancements
  - Oracle XDK

11gR1 : In Place Schema Evolution

- Allows simple changes to registered XML schemas with zero down-time
  - No data copy required
  - Schema change takes a few seconds regardless of amount of data.
- Changes to the XML must not invalidate existing documents.
  - Add optional elements and attributes
  - Adding new values to enumerations
  - Increase in length

11gR1 XMLType partitioning

- Leverages the 11gR1 “Ref-Based” partitioning feature
- Nested Tables are partitioned using the same key as the top level XMLType table
  - Includes supports for “out-of-line” storage models
- Improved Manageability
  - “partition maintenance” operations
  - All operations are performed on the XMLType table
- Improved query optimization
  - “partition pruning”

11gR1 : Intelligent Defaults

- Avoids most common limitations and bottlenecks
  - Default storage model is now collections stored as nested-tables
- Default organization is heap-organized nested-tables
  - Optimal re-write for XQuery expressions and XMLTable-based operations
  - Partitioning and indexing
  - Text based indexing
- Optional, automatic horizontal partitioning for wide structures

11gR1 General Improvements

- XMLCast and XMLExists operators
  - XMLCast : XQuery for fragment extraction
  - XMLExists : XQuery in where clause
- Streaming XSLT
  - Significant performance improvements when generating large documents
  - XSLT output no-longer restricted to well-formed XML
- Large node handling eliminates current 64K limit on size of a text-node.
- Stream based replication support for Text-based XML storage

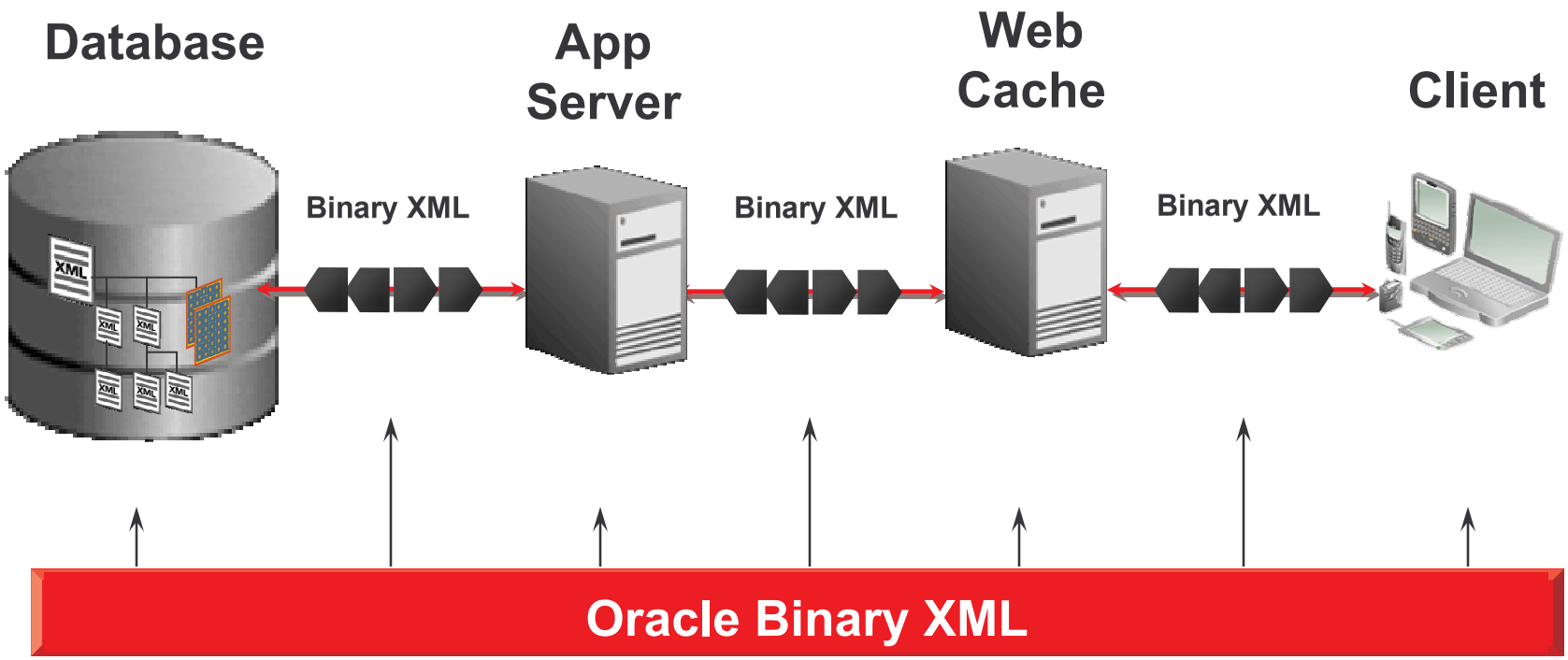
11gR1 Performance Improvements

- Significant performance improvements
  - Up to 10x performance improvement when ingesting XML
  - Up to 10x performance improvement when generating XML
    - XMLAgg optimization
  - General improved optimization of XQuery expression on Schema-Optimized storage
  - Significant optimization for fragment level DML operations on collections

Agenda

- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
  - Improvements for Schema-Optimized XML
  - Schema-Less XML
  - XML Index
  - Database-Native Web Services
  - Repository Improvements
  - Security Enhancements
  - Oracle XDK

11gR1 : Binary XML



11gR1 : Binary XML Storage Option

- Supports Schema-Less and Schema-based XML
- Post Parse representation of XML
- Single format for “on-disk”, “in-memory” and “on-the-wire”
- Optimized for indexing and fragment extraction
- Flexible XML Schema Support
  - No “Schema-Evolution” issues
  - Reduces storage requirements
  - Heterogeneous XML documents in a single table or column
- Tight integration with Oracle Secure Files

11gR1 Binary XML advantages

- Compact Storage Format
  - Tag-names are tokenized
  - Text-nodes and attribute values are stored in native representation, rather than text format
- Reduced CPU and Memory overhead
  - Pull parser, streaming validation and streaming XPath avoid use of DOM Tree
  - Single format avoids parse and serialize issue when moving XML between application tiers
- Reduced Network Overhead
  - Post-parse format used when XML moves between tiers

11gR1 Binary XML advantages

- High performance fragment access and extraction
  - Streaming XPath allows multiple nodes to be accessed in a single operation
- High performance update
  - Sliding inserts
  - Partial re-validation following update
- Locale Sensitive
  - Support XLiff, allowing locale-sensitive retrieval of content

Agenda

- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
  - Improvements for Schema-Optimized XML
  - Schema-Less XML
  - XML Index
  - Database-Native Web Services
  - Repository Improvements
  - Security Enhancements
  - Oracle XDK

11gR1 : XML Index

- New universal index for Binary and Text based XMLType storage models
- Addresses all known issues with CTX-XPath index
- Optimizes most common classes of Path Expressions
  - Recursive, Relative, Lazy (//)
- Accelerates path & value based predicates
- Fully type aware
  - Optimizes numeric and date range predicates
- Fully namespace aware

11gR1 : Totally transparent

- Support fragment extraction as well as fragment existence
- Supports all XQuery operators
  - XMLQuery, XMLTable, XMLExists, XMLCast
- Supports DML operators and legacy XPath operators
  - updateXML, insertChildXML,
  - existsNode, extractValue, xmlSequence …
- No code changes required
- Certain queries are executed totally against the index

11gR1 : XML Index Optimizations

- XML Index can get large / expensive to maintain
  - By default all possible paths are indexed by default
  - High disk usage
  - Impact on throughput
- XML index optimization
  - Asynchronous Operation
  - Path Sub-setting
  - Partial re-indexing

11gR1 : XML Index Optimization

- Synchronous and Asynchronous indexes
  - Determined at Index creation time
- Synchronous Mode
  - Insert operation does not complete until indexing complete
  - Query results always include DML changes
- Asynchronous Mode
  - Insert operation does not wait for indexing to take place
  - Query results may not reflect latest DML changes
  - Index maintain takes place automatically in near-real time
- No Fragmentation Issues in either mode

XML Index Optimizations

- Path Subsetting
  - Provides control over which nodes within the document are indexed
  - Trade index size, throughput against query performance
  - Index can only optimizes performance for the indexed Paths
  - Can add or remove paths dynamically
- Partial re-indexing
  - Only re-index modified content
  - Requires Binary XML and Secure Files

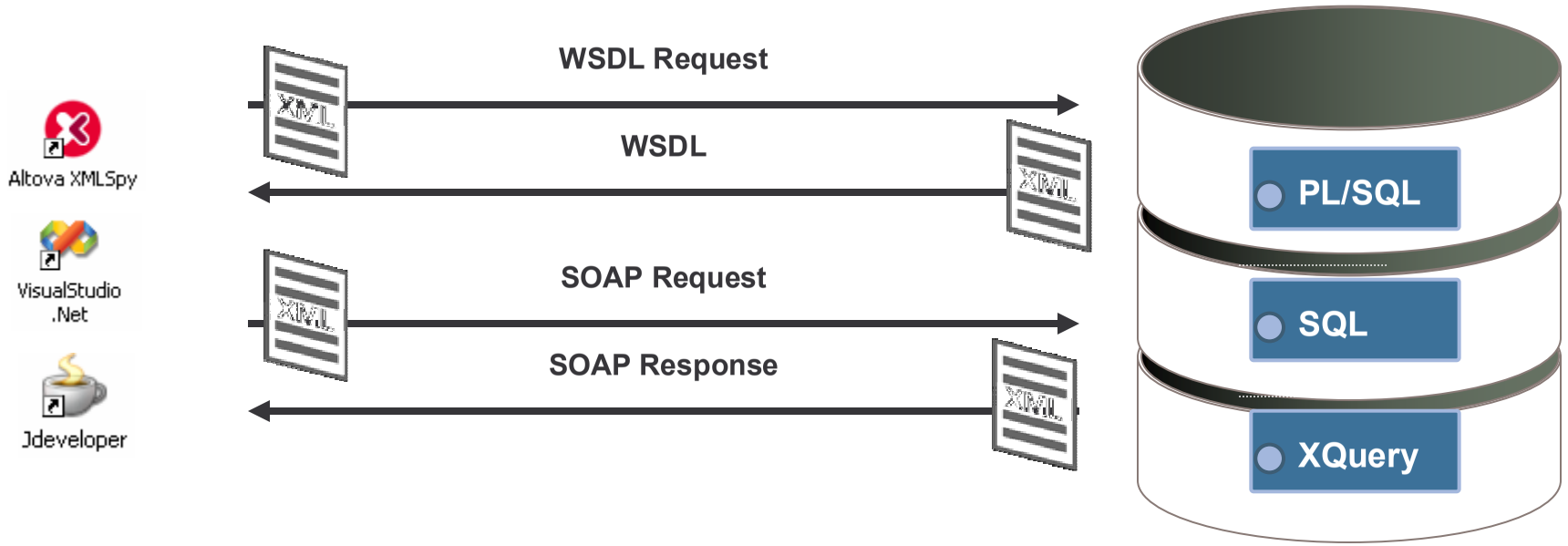
11gR1 : XML Index features

- Repository integration
  - Create XML Index on XML DB repository
  - Optimize queries on meta-data and content
- Hybrid Storage
  - Create XML index on fragments in Schema-Optimized storage that are mapped to CLOB
  - Enhanced query capabilities
- ‘Data Guide’
  - Index can provide meta-data about the XML it has indexed
- Tight integration with Oracle Text
  - Index driven full-text search

Agenda

- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
  - Improvements for Schema-Optimized XML
  - Schema-Less XML
  - XML Index
  - Database-native Web Services
  - Repository Improvements
  - Security Enhancements
  - Oracle XDK

Database-native Web Services



Database-native Web Services

- ‘Zero-Development’, ‘Zero-Deployment’ solution for publishing PL/SQL packages.
  - Any package method, function or procedure can be accessed as a SOAP end-point
- Leverages the Oracle XML DB HTTP Server
  - No additional infrastructure required
- Automatic generation of WSDL
  - URL to Package, Function or Procedure mapping scheme
- Uses XML DB infrastructure for processing request and generating response
- Includes ‘SQL Query’ and ‘XQuery’ Services

Agenda

- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
  - Improvements for Schema-Optimized XML
  - Schema-Less XML
  - XML Index
  - Database-Native Web Services
  - Repository Improvements
  - Security Enhancements
  - Oracle XDK

XML DB repository events

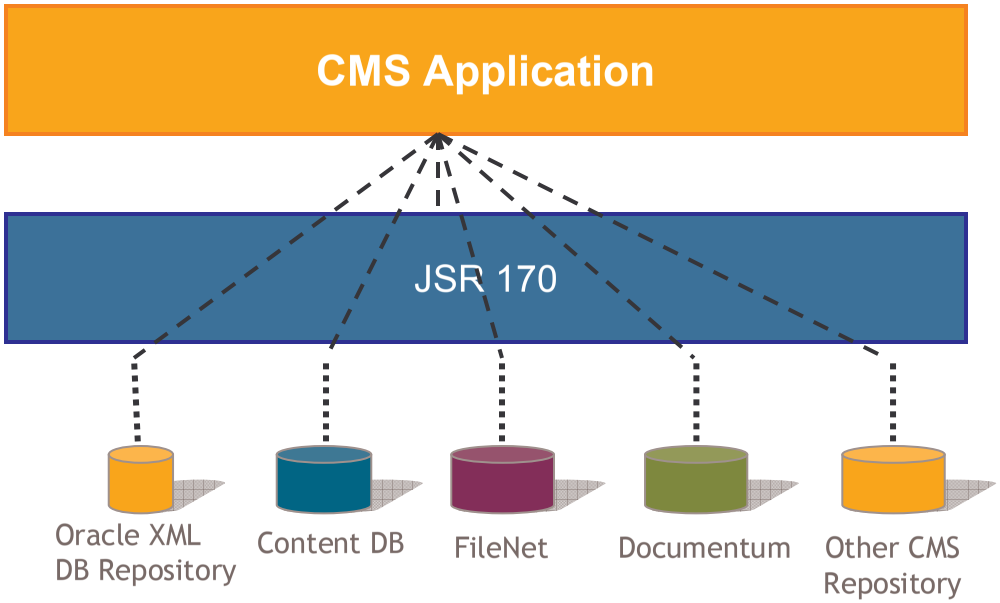
- ‘Triggers’ for the XML DB repository.
  - Enables ‘server-enforced’ rules for XML DB repository
  - PL/SQL procedures automatically executed as a result of repository activity
- Multiple-models
  - Repository Wide
  - Folder or Folder-Tree specific
  - Content specific
- Associate code with
  - CRUD (Create, Replace, Update and Delete) operations on documents
  - Folder Link and Unlink

XML DB repository enhancements

- XLink and XInclude
  - Support for compound documents
  - XML centric referential integrity
  - Enforcement of XLink based references
- Extended Links
  - Support for ‘Soft’ links

XML DB Content Connector: JSR-170

All Enterprise Content Repository



- Standardized APIs accessing XML DB content repository
  - Manage structure, unstructured and semi-structured content in one repository
  - Traverse, query, access, and operate data using standard APIs
- Support JSR 170 Level 1 and Level 2

Agenda

- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
  - Improvements for Schema-Optimized XML
  - Schema-Less XML
  - XML Index
  - Database-Native Web Services
  - Repository Improvements
  - Security Enhancements
  - Oracle XDK

XML DB Security enhancements

- Support for DAV ACL
  - Security interoperability with emerging DAV Clients
- ACL inheritance
  - Define common security model
    - Organization-wide policies
    - Content-specific policies
- User-defined ACLs
  - Use ACL mechanism to manage access to non XML DB objects

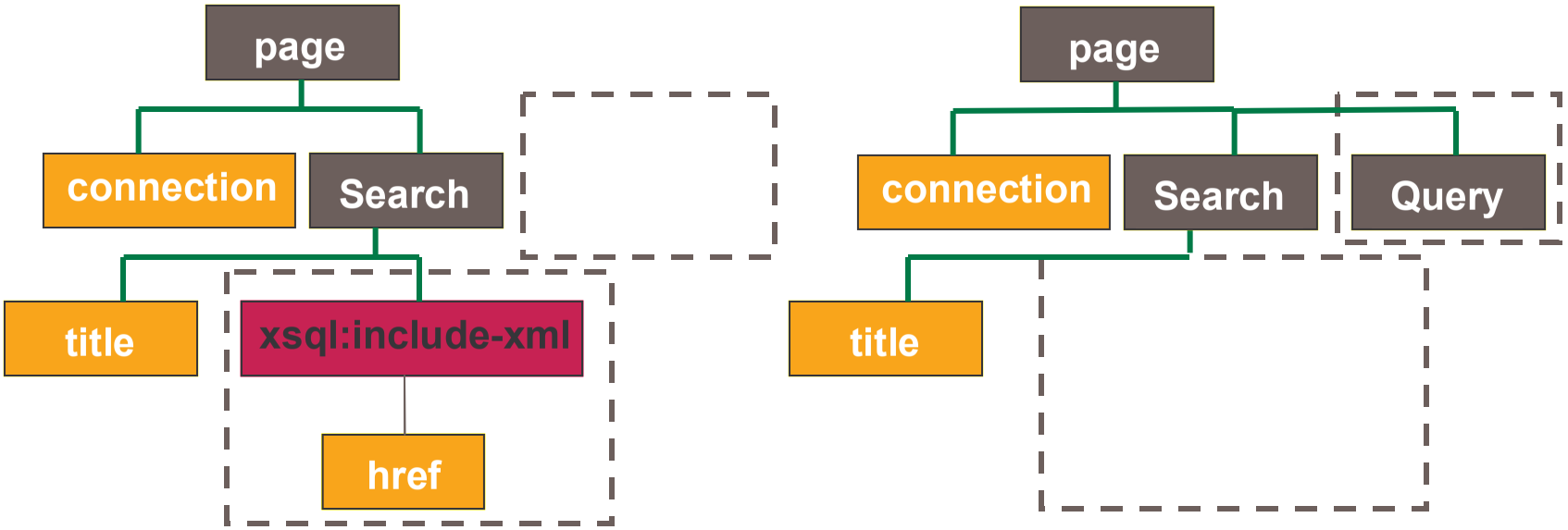
Agenda

- XML Trends
- Introducing Oracle XML DB
- What's new in Oracle XML DB 11g
  - Improvements for Schema-Optimized XML
  - Schema-Less XML
  - XML Index
  - Database-Native Web Services
  - Repository Improvements
  - Security Enhancements
  - Oracle XDK

XDK Enhancements

- XDK “C”
  - Pull Parser
  - Streaming Validator
  - Binary XML integration
- XDK “J”
  - Binary XML integration
  - Scaleable ‘DOM’
- XML Diff and XMLPatch
  - Stand alone utility
  - ‘C’ API
  - PL/SQL API
  - Java ‘forthcoming’

XML Diff and Patch



- Compare two XML documents
- Represent the difference in XML
- Apply and control the changes

XMLDiff Output

```
<xd:xdiff xsi:schemaLocation="http://xmlns.oracle.com/xdb/xdiff.xsd
                              http://xmlns.oracle.com/xdb/xdiff.xsd"
   xmlns:xs="http://www.w3.org/2001/XMLSchema"
   xmlns:xdb="http://xmlns.oracle.com/xdb"
   xmlns:xd="http://xmlns.oracle.com/xdb/xdiff.xsd"
   xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance">
 <?oracle-xmldiff operations-in-docorder="true" output-model="snapshot" diff-
   algorithm="global"?>
   <xd:append-node xd:node-type="element" xd:parent-xpath=
                   "/xs:schema[1]/xs:simpleType[6]/xs:restriction[1]">
      <xd:content>
        <xs:enumeration value="E999" />
      </xd:content>
    </xd:append-node>
</xd:xdiff>
```

Q&A

ORACLE